

### REMARKS

Claims 1-5, and 7-23 are presented for examination. Claims 1 and 7 are currently amended. No claim is cancelled.

Independent claims 1, 7, 10, and 11 (and their corresponding dependent claims) were rejected under 35 U.S.C. § 101 as not being in the technological arts. Specially, the claims are noted as being so broad as to encompass a pen and paper and a user accomplishing the claim. Claims 1 and 7 are amended to specify that the recited method is for use in the categorizing of documents "in an electronic system", so as to overcome the present rejection. In regards to claims 10 and 11, Applicants respectfully point out that claims 10 and 11 already recite, "A document categorizing **apparatus**", are thus believed to already recite a limiting structure.

Claims 1-5, 10, 12, 14, and 16-18 are rejected under 35 U.S.C. § 103(a) as being unpatentable over Zamir et al. in view of Davies et al. (U.S. Pat. 5,931,907). Claims 7-9, 11, 13, 15, and 20-23 are similarly rejected under 35 U.S.C. § 103(a) over Zamir and Davies, and further in view of Wu (U.S. Pat. 5,991,756).

In regards to claims 1-5, 10, 12, 14, and 16-18, the Office Action concedes that,

"Zamir fails to teach that said cluster merging process defines said degree of relation between multiple clusters under consideration as the number of distinct files common to all of said clusters under consideration multiplied by a predefined multiplication factor divided by a total sum of all the files in said clusters under consideration".

However, the Office Action then cites col. 8, lines 5-8 of the Davies reference to support an assertion that Davies shows this limitation. Specifically, the Office Action explains that Davies shows that for each pair of documents (Di and Dj), a coefficient is calculated using formula:  $2*[Di.ANDGATE.Dj]/[Di]+[Dj]$ ".

Applicants respectfully point out that Col. 8, lines 5-8 of Davies explain that this formula is used to compare two individual documents in order to construct a cluster of similar documents. That is, the Davies references does not

teach or suggest the comparing/combining of clusters, but rather shows the comparing of individual pairs of documents prior to the formation of clusters.

In other words, the Office Action appears to be equating a single document to a cluster of documents (and equating individual words within a document to individual documents within a cluster). The cited formula above clearly states that it is comparing the occurrences of words within two documents,  $D_i$  and  $D_j$ . This formula does not indicate if it can (or indeed how it could) be used to compare multiple clusters, and how to rate the similarity between the multiple clusters.

Furthermore, the language of claim 1 clearly states that the degree of relation between the multiple clusters is the number of distinct files common to all clusters under consideration multiplied by a predefined multiplication factor and divided by the total sum of all files in all the clusters under consideration. The formula above cited in the Office Action (Davies, Col. 8, line 13) states that the numerator is divided by  $D_i$ , and the result is summed to  $D_j$ . This clearly is not the sum of all files in all clusters under consideration. Thus, even if the teachings of Davies were combined with those of Zamir, one would not achieve the present invention.

Additionally, it must be emphasized that Davies shows a specific formula (as recited above) for comparing pairs of documents for the purpose of constructing clusters. Zamir similarly shows a method for comparing documents for the purpose of creating clusters. Thus, if the teaching of Davies were combined with those of Zamir, one would expect Davies formula for comparing documents to replace Zamir's method of comparing documents. There is teachings, suggestion, or incentive to attempt to apply Davies' formula for comparing documents to Zamir's method for comparing clusters, especially since no teaching is provided for explaining how Davies' formula might be altered to be used in rating similarities between clusters.

The Office Action's equating of a document (as taught by Davies) to a cluster of similar documents (as recited in the present claims) is the basis for many of its other rejections. For example in reference to claim 2, the Office Action states that Davies shows the combining of two "clusters". As explained

above, this is an error. Davies shows the rating of similarities between two "documents", not clusters.

In reference to claim 7 (and similarly claims 11 and 13) the Office Action states that U.S. Pat. 5,991,756 to Wu, teaches in Fig. 5,

"the display of a search result including category names listed in a hierarchical format, which are grouped to links of similar documents. Though Wu does not call these categories/sub-categories name clusters, the fact that each link in the hierarchy from left to right ...represents a group of similar documents, by definition can be thought of as clusters of similar documents. As one traverses the hierarchy from left to right, one traverses the cluster hierarchy from general to more specific. This traversal also inherently represents a degree of similar documents. Though not specifically taught by Wu, it would have been obvious to one of ordinary skill in the art at the time the invention to conclude that such a portrayal of document cluster names as seen in fig. 5 constitutes the claimed first and second listing formats based on interpretation of similarity measures".

Here, the Office Action appears to again be overlooking the difference between a cluster of documents, and a single document. The Office Action also appears to be misunderstanding the meaning of the term "name format". A name format is predefined naming convention applied to files, such as, for example; (1) requiring that a directory path be made part of a file name; (2) requiring that files be given a common suffix following a period (such as "\*.doc", or "\*.pdf", or "\*.zip", etc.); (3) require that the file name include the path in a file-structure tree under which the file is listed; (4) or any other arbitrarily defined pattern. The point is that the name is based on a predefined pattern. As noted above, many files can have different names, but follow a single "name format".

Firstly in Fig. 5, Wu shows document names on a local machine (54) and document names at specific web sites (58). These are names of individual documents, not clusters of documents. Thus, there is no inherent tendency of "general to more specific" when reading a document name from left to right because it is a single document name, not a series of links between collections of documents. It just happens that the file names follow a common naming format of listing the sub-categories under which the document is located.

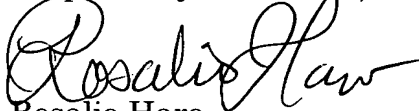
That is, both documents types 54 and 58 indicated in the Office Action follow ONE naming format: i.e. the document titles define a category/sub-

category path to the target document (Col. 8, lines 58-60). This naming format (which requires the listing the categories traversed to reach a specific document) is the same for document types 54 and 58, and thus represents only ONE naming format.

Conversely, claim 7 not only requires TWO naming formats, but also requires that the decision of which of the two naming formats is used depends upon the results of similarity calculations. Wu does not teach or suggest two naming formats, nor does it teach or suggest that one naming format takes precedence over another based on a similarity calculation.

In view of the foregoing amendments and remarks, Applicants respectfully request favorable reconsideration of the present application.

Respectfully submitted,



Rosalio Haro

Registration No. 42,633

Please address all correspondence to:

Epson Research and Development, Inc.  
Intellectual Property Department  
150 River Oaks Parkway, Suite 225  
San Jose, CA 95134  
Phone: (408) 952-6000  
Facsimile: (408) 954-9058  
Customer No. 20178

Date: April 21, 2005